



特開平10-187358

(43) 公開日 平成10年(1998) 7月14日

(51) Int. Cl.<sup>6</sup>  
G 0 6 F 3/06  
11/30

識別記号  
5 4 0  
3 0 4

F I  
G 0 6 F 3/06 5 4 0  
3 0 4 N  
11/30 E

審査請求 未請求 請求項の数 5 ○ L (全 8 頁)

(21) 出願番号 特願平8-344998  
(22) 出願日 平成8年(1996)12月25日

(71) 出願人 000005108  
株式会社日立製作所  
東京都千代田区神田駿河台四丁目6番地  
(72) 発明者 八木沢 育哉  
神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内  
(72) 発明者 大枝 高  
神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内  
(72) 発明者 松並 直人  
神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内  
(74) 代理人 弁理士 小川 勝男

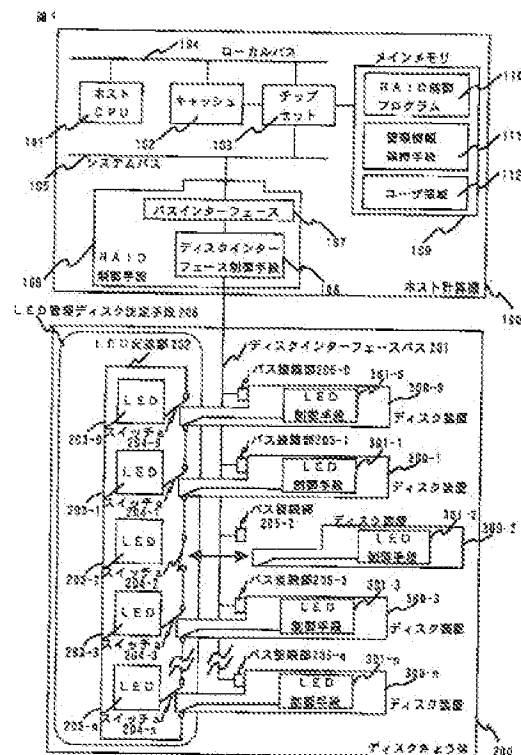
最終頁に続く

(54) 【発明の名称】 ディスクアレイ装置

(57) 【要約】

【課題】故障したディスク装置を特定するLEDを他のディスク装置を用いて点灯させることで、RAIDの制御を行う手段とディスクきょう体を接続する目的で特別な専用線を用いることなく、低価格でディスクきょう体を管理する方法を提供することができなかった。

【解決手段】ディスク装置には、他のディスク装置のLED発光を制御するLED制御手段を設け、ディスクきょう体には、複数のディスク装置の各々に対応するLEDと、あるディスク装置とそれに対応するLEDを発光する制御を行うディスク装置を一意に決定するLED管理ディスク決定手段を設け、ホスト計算機のRAID制御プログラムには、あるディスク装置とそれに対応するLEDを発光する制御をおこなうディスク装置に関連づけるID管理テーブルを作成、参照する故障検出LED制御手段を設け、ホスト計算機には、上記のID管理テーブルを格納する管理情報保持手段を設けたことにより達成することができる。



## 【特許請求の範囲】

【請求項1】ホスト計算機がRAID制御手段とRAID制御プログラムを備え、ディスクきょう体に複数台のディスク装置を格納し、ホスト計算機とディスクきょう体間をディスクとのデータ転送を行うディスクインターフェースバスで接続した計算機システムにおいて、ディスク装置には、他のディスク装置のLED発光を制御するLED制御手段を設け、

ディスクきょう体には、複数のディスク装置の各々に対応するLEDと、あるディスク装置とそれに対応するLEDを発光する制御を行うディスク装置を一意に決定するLED管理ディスク決定手段を設け、ホスト計算機のRAID制御プログラムには、あるディスク装置とそれに対応するLEDを発光する制御を行うディスク装置に関連づけるID管理テーブルを作成、参照する故障検出LED制御手段を設け、ホスト計算機には、上記のID管理テーブルを格納する管理情報保持手段を設けたことを特徴とする計算機システム。

【請求項2】ホスト計算機がRAID制御手段とRAID制御プログラムを備え、ディスクきょう体に複数台のディスク装置を格納し、ホスト計算機とディスクきょう体間をディスクとのデータ転送を行うディスクインターフェースバスで接続した計算機システムにおいて、ディスク装置には、他のディスク装置とインターフェースバスの接続状況を監視、記録する接続状況記録手段を設け、ディスクきょう体には、複数のディスク装置の各々がインターフェースバスと接続されているかどうかを示す接続調査手段と、あるディスク装置とそのディスク装置がインターフェースバスと接続されているかどうかを監視するディスク装置を一意に決定する活線挿抜管理ディスク決定手段を設け、ホスト計算機のRAID制御プログラムには、あるディスク装置とそのディスク装置がインターフェースバスと接続されているかどうかを監視するディスク装置に関連づけるID管理テーブルを作成、参照する活線挿抜制御手段を設け、ホスト計算機には、上記のID管理テーブルを格納する管理情報保持手段を設けたことを特徴とする計算機システム。

【請求項3】上記請求項1、2記載の計算機システムにおいて、RAID制御プログラムはホスト計算機のホストCPUにより実行されることを特徴とする計算機システム。

【請求項4】上記請求項1、2記載の計算機システムにおいて、RAID制御手段はRAIDの制御/管理専用のプロセッサを有し、RAID制御プログラムはRAIDの制御/管理専用のプロセッサにより実行されることを特徴とする計算機システム。

【請求項5】上記請求項1、2記載の計算機システムにおいて、RAID制御手段およびRAID制御プログラムはホスト計算機の外部にあるRAID装置にあり、上記RAID装置とホスト計算機とがデータ転送経路で接

続されていることを特徴とする計算機システム。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は主として、コンピュータシステムの記憶システムとして用いられているディスクアレイシステムのための制御方式に係わり、特に、ディスク装置格納きょう体の管理手法に関するものである。

## 【0002】

【従来の技術】ディスクアレイシステムは、RAID(Redundant Arrays of Inexpensive Disks)と呼ばれ、複数のディスク装置をアレイ状に配置した構成をとり、データの読み出し・書き込み要求をディスクの並列動作によって高速に処理するとともに、データにパリティ(冗長)データを付加することによって信頼性を向上させた記憶装置である。

【0003】RAIDを記憶装置として利用するホスト計算機は、連続運転を要求される用途に使用されることも多く、したがって、RAID内のディスク装置が故障した場合は、ホスト計算機やRAIDの電源を切らずにディスク装置の交換を行う必要性が高い。このように電源を入れた状態で行うディスク装置の交換は、活線挿抜と呼ばれ、ディスク装置が格納されているディスクきょう体で実現されている。このディスクきょう体は、ホスト計算機と共通の場合もあるし、ホスト計算機とは別の場合もある。

【0004】ディスクきょう体には、交換すべき故障ディスク装置がどれかをユーザに視覚的に知らせるためのLED等の表示手段や、ユーザにより活線挿抜が行われたことを検出するための手段が設けられている。

【0005】ディスクきょう体管理の従来技術として、米CONNER社が発行した「SCSI Accessed Fault-Tolerant (SAF-TE) Enclosures White Paper」がある。SAF-TEは、ディスクの状態管理をディスクきょう体の中にある管理専用の1つのSCSIデバイスで実現したものである。このSAF-TEは、故障したディスク装置をユーザに示すのにLEDを用いており、ホストからのSCSIコマンドによる要求に応じて、管理専用SCSIデバイスがLEDの点灯/消灯制御を行っている。

## 【0006】

【発明が解決しようとする課題】従来の技術によれば、ディスクの状態管理をSCSIバスに接続した1つの管理専用SCSIデバイスで実現しているので、ディスク装置を用いてRAIDを構築する際には、RAIDの制御を行う手段とディスクきょう体を接続するのに特別な専用線を必要としないという効果がある。

【0007】しかし、SCSIデバイスには、制御用のプロトコル解析プロセッサと、専用のインターフェース

コントローラが必要であるため、高価格になってしまうという問題がある。

【0008】さらに、プロトコル解析プロセッサ制御用のプログラムは、開発工数が多いため、その結果高価格なシステムになってしまうという問題がある。

【0009】さらに、ディスクきょう体に複数のSCSIを内蔵した場合、1台のディスクきょう体を管理するためにSCSI本数分の管理専用SCSIデバイスが必要になり、高価格なシステムになってしまうという問題がある。

【0010】本発明の第1の目的は、ディスク管理専用デバイスを使わず、かつ、RAID制御を行う手段とディスクきょう体を特別な専用線をを用いないで接続するとともに、故障ディスク装置を示すLEDの点灯制御を低価格で提供することである。

【0011】本発明の第2の目的は、ディスク管理専用デバイスを使わず、かつ、RAID制御を行う手段とディスクきょう体を特別な専用線をを用いないで接続するとともに、ディスク装置交換時の活線挿抜制御を低価格で提供することである。

【0012】

【課題を解決するための手段】前記第1の目的を達成するために本発明は、ホスト計算機にRAID制御手段とRAID制御プログラムを設け、ディスクきょう体に複数台のディスク装置を格納し、ホスト計算機とディスクきょう体間をディスクとのデータ転送を行うディスクインターフェースバスで接続した計算機システムであり、ディスク装置には、他のディスク装置のLED発光を制御するLED制御手段を設け、ディスクきょう体には、複数のディスク装置の各々に対応するLEDと、あるディスク装置とそれに対応するLEDを発光する制御を行うディスク装置を一意に決定するLED管理ディスク決定手段を設け、ホスト計算機のRAID制御プログラムには、あるディスク装置とそれに対応するLEDを発光する制御を行うディスク装置に関連づけるID管理テーブルを作成、参照する故障検出LED制御手段を設け、ホスト計算機には、上記のID管理テーブルを格納する管理情報保持手段を設ける。

【0013】また、前記第2の目的を達成するために本発明は、ホスト計算機にRAID制御手段とRAID制御プログラムを設け、ディスクきょう体に複数台のディスク装置を格納し、ホスト計算機とディスクきょう体間をディスクとのデータ転送を行うディスクインターフェースバスで接続した計算機システムであり、ディスク装置には、他のディスク装置とインターフェースバスの接続状況を監視、記録する接続状況記録手段を設け、ディスクきょう体には、複数のディスク装置の各々がインターフェースバスと接続されているかどうかを示す接続調査手段と、あるディスク装置とそのディスク装置がインターフェースバスと接続されているかどうかを監視する

ディスク装置を一意に決定する活線挿抜管理ディスク決定手段を設け、ホスト計算機のRAID制御プログラムには、あるディスク装置とそのディスク装置がインターフェースバスと接続されているかどうかを監視するディスク装置に関連づけるID管理テーブルを作成、参照する活線挿抜制御手段を設け、ホスト計算機には、上記のID管理テーブルを格納する管理情報保持手段を設ける。

【0014】

【発明の実施の形態】本発明の第1の実施形態を図を用いて説明する。第1の実施形態は、近傍のディスク装置でLEDを発光することによって、故障したディスク装置をユーザに通知するためのものである。

【0015】図1は本発明の第1実施形態のシステム構成図である。図1において、100はホスト計算機、200はディスク装置を格納するディスクきょう体、300-0~300-nはディスク装置である。

【0016】ホスト計算機100は、ホストCPU（中央処理演算機）101と、高速リード・ライト可能なメモリであるキャッシュ102と、データ転送制御を行うチップセット103を備え、これらをローカルバス104で接続している。チップセット103には、データやプログラムを格納するメインメモリ109と、システムバス105を接続している。メインメモリ109は、後述するRAID制御プログラム110と、図3に示すようなID管理テーブル130を有する管理情報保持手段111と、ユーザデータを格納するユーザ領域112を保有している。

【0017】第1の実施形態において、RAID制御手段106は、ホスト計算機100に内蔵する拡張ボード型である。RAID制御手段106は、バスインターフェース107とディスクインターフェース制御手段108を備え、バスインターフェース107を介して、システムバス105に接続する。ディスクインターフェース制御手段108とバスインターフェース107は接続しており、ディスクインターフェース制御手段108は、ディスクインターフェースバス201を介してディスクきょう体200と接続している。第1の実施形態のRAID制御手段106は、RAID管理、制御のためのプロセッサを持たず、メインメモリ109にあるRAID制御プログラム110内のRAID制御プログラムをホストCPU101で実行することでRAIDの管理、制御を行う。

【0018】図2は第1の実施形態のRAID制御プログラム110のブロック図である。RAID制御プログラム110は、アプリケーションからの要求を受信する要求受信部120と、アプリケーションに要求終了を通知する要求終了部121と、RAIDの管理、制御を行うRAID制御部122と、ディスク装置300-0~300-nへのコマンド発行と終了通知を受信するディ

ディスクインターフェース制御手段制御部123と、ディスクきょう体200内部の故障したディスク装置をユーザへ通知するためにどのLEDを発光すべきかを管理する故障検出LED制御手段124からなる。

【0019】図1に示す、ディスクきょう体200は、ディスクインターフェースバス201と、バス接続部205-0～205-nと、スイッチa204-0～204-nで間を接続したLED配線部202と、それぞれがスイッチa204-0～204-nの片方の接続先であるLED（発光ダイオード）203-0～203-nからなる。スイッチa204-0～204-nとLED配線部202を総称してLED管理ディスク決定手段206と呼ぶこととし、LED管理ディスク決定手段206は後述するように、故障したディスク装置301と、そのディスク装置301に対応するLED203を発光する管理ディスクの関係を一意に決定する。スイッチa204-0～204-nのもう片方の接続先は、LED配線部202を介したとなりのスイッチaである。なお、nはディスクインターフェースバス201が接続することができるディスク装置の最大数から1を引いた数以下の自然数である。バス接続部205-0～205-nは、ディスクインターフェースバス201に接続している。

【0020】ディスク装置300-0～300-nは、ディスクきょう体200に格納するもので、それぞれの内部には、LED発光コマンドを受信し、LEDの発光制御を行うLED制御手段301-0～301-nを有する。

【0021】各ディスク装置300-0～300-nをディスクきょう体200に格納していない状態では、対応する各スイッチa204-0～204-nは、LED配線部202をつないで、となりのスイッチaと接続する状態になっている。図1では、スイッチa204-2が、この状態となっている。

【0022】各ディスク装置300-0～300-nをディスクきょう体200に格納すると、対応する各バス接続部205-0～205-nと接続され、ホスト計算機100とデータ、制御情報等のやりとりができるようになる。各ディスク装置300-0～300-nが、対応する各スイッチa204-0～204-nをLED203-0～203-nに接続するように切り換える。図1では、スイッチa204-0などがこの状態となっている。また、この状態のときには、各ディスク装置300-0～300-n内のLED制御手段301-0～301-nが、対応する各スイッチa204-0～204-nと図1のように接続されることで、となりのスイッチaに接続される。

【0023】したがって、図1のディスク装置300-0とディスク装置300-1のように、2台のディスク装置300が連続して格納されている場合は、ディスク

装置300-0内部のLED制御手段301-0は、スイッチa204-1を介して、となりのディスク装置300-1に対応するLED203-1に接続されている状態となる。

【0024】また、図1のディスク装置300-1とディスク装置300-3のように、間のディスク装置が格納されていない状態で2台のディスク装置300が格納されている場合は、ディスク装置300-1内部のLED制御手段301-1は、スイッチa204-2とスイッチa204-3を介して、2台となりのディスク装置300-3に対応するLED203-3に接続されている状態となる。

【0025】同様にして、2台以上の間隔が空いている場合も、ディスク装置が格納されていない場所のスイッチa204がLED配線部202をつなぐように接続されているので、あるディスク装置300内部のLED制御手段301は、LED制御手段301がLED配線部202とつながる側で配線的に一番近くに格納されているディスク装置に対応するスイッチa204と接続される。

【0026】上記のようにしてLED制御手段301に接続されたLED203は、LED制御手段301によって発光させることができるものとする。

【0027】LED配線部202は、スイッチa204-0～204-nによって個々につながれており、図1において一番端にあるスイッチa204-nは、スイッチa204-0とつながれているので、どのディスク装置300-0～300-nをディスクきょう体200に格納しても、となりのスイッチa204との接続条件は同じになる。

【0028】なお、LED203-0～203-nは、末尾番号が同じディスク装置300-0～300-nと対応していることが分かるような配置がしてあればよい。また、第1の実施形態では、説明の便宜上、ディスクきょう体200内のディスクインターフェースバス201上での各ディスク装置300のID（識別番号）を各ディスク装置300-0～300-nの末尾番号と同じであるとする。

【0029】次に、どのディスク装置300が故障したかをユーザに知らせるための動作を示す。第1の実施形態においては、故障したディスク装置300に対応するLED203を別のディスク装置300が発光することで交換すべき故障したディスク装置300をユーザに知らせる。図1の例で言えば、ディスク装置300-1に対応するのは、末尾番号が同じLED203-1である。第1の実施形態において、LED203を発光させるディスク装置300を管理ディスクと呼び、そのIDを管理IDと呼ぶこととする。また、この管理ディスクによって故障したことをユーザに知らされるディスク装置300を被管理ディスクと呼び、そのIDを被管理ID

Dと呼ぶこととする。ディスクきょう体200に格納された各ディスク装置300は、管理ディスクにも被管理ディスクにもなる。

【0030】第1の実施形態においては、上記のシステム構成で述べたような、LED配線部202とスイッチa204を介して接続されているLED制御手段301とLED203の関係で管理ディスクと被管理ディスクが決定される。接続されているLED制御手段301を有するディスク装置300が管理ディスクとなり、LED203に対応したディスク装置300が被管理ディスクとなる。図1の例で言えば、ディスク装置300-0が管理ディスクだとすると、LED203-1に対応するディスク装置300-1が被管理ディスクである。管理ディスクと被管理ディスクをこのように決定することで、管理ディスクによって、被管理ディスクに対応するLED203を発光することができる。

【0031】まず、管理ディスクと被管理ディスクに関連づけるID管理テーブル130の作成方法について述べる。システムの初期化時に、RAID制御部122はディスクインターフェース制御手段制御部123を介して、ディスクインターフェースバス201上のすべてのIDに対して、実際にディスク装置が格納されているかどうかを確認するコマンドを発行する。この結果をもとに、故障検出LED制御手段124は、ID管理テーブル130を作成する。ディスクインターフェースバス201上に存在できる最大ID数nは、システムを構成する段階で一意に決めておけばよい。ディスク装置が実際に格納されていたIDだけが管理IDおよび被管理IDになることができ、その中で最も小さいIDに対して最も大きいIDの番号を管理IDとして割り当てる。また、ディスク装置が格納されていたその他のIDに対しては、そのIDより大きく、最も番号に近いIDを管理IDとして割り当てる。図1のようなシステム構成の場合の例を図3のID管理テーブル130に示す。故障検出LED制御手段124は、このようにしてID管理テーブル130を作成し、管理情報保持手段111に保持する。

【0032】次に、どのディスク装置300が故障したかをユーザに知らせるためのLEDの発光方法を示す。あるディスク装置300が故障すると、ディスクインターフェース制御手段108を介して、ディスクインターフェース制御手段制御部123にどのディスク装置300が故障したかの通知が送られる。この情報をもとにして、RAID制御部122は、故障検出LED制御手段124に発光すべきLED203を問合わせ、故障検出LED制御手段124は、ID管理テーブル130をもとに故障したディスク装置300に関連する管理IDをRAID制御部122に知らせる。その管理IDのディスク装置300に対して、RAID制御部122は、ディスクインターフェース制御手段制御部123を介

し、LED発光コマンドを発行する。LED発光コマンドを受け取った管理ディスクのLED制御手段301は、LED配線部202とスイッチa204を介して接続されている被管理ディスクに対応したLED203を発光させる。

【0033】このようにして、管理ディスクは被管理ディスクに対応したLED203を発光し、ユーザは、この発光したLEDによって故障したディスク装置300を特定することができる。また、故障したディスク装置300を抜くときに、対応するLED203とスイッチa204が離れるので、この状態をLED制御手段301で認識することで、その後にLED203を再び発光するのを防ぐことができる。

【0034】以上の構成、動作により、本発明によれば、ディスク管理専用デバイスを使わず、かつ、RAID制御を行う手段とディスクきょう体を特別な専用線を用いなくて接続するとともに、故障ディスク装置を示すLEDの点灯制御を低価格で提供することができるという効果がある。

【0035】本発明の第2の実施形態を図を用いて説明する。第2の実施形態は、近傍のディスク装置から故障したディスク装置のIDを監視することにより、故障したディスク装置がユーザによって抜かれたことと、新しいディスク装置がユーザによって挿入されたことをホスト計算機に通知するためのものである。

【0036】図4は本発明の第2実施形態のシステム構成図である。図1との違いだけを述べる。ディスクきょう体200において、スイッチa204-0~204-nの代わりにそれぞれスイッチb208-0~208-nが、LED配線部202の代わりに接続調査配線部209がそれぞれ設けられているが、その動作と接続形態は図1と同じである。スイッチb208-0~208-nと接続調査配線部209を総称して活線挿抜管理ディスク決定手段210と呼ぶこととし、活線挿抜管理ディスク決定手段210は故障したディスク装置301と、そのディスク装置301の活線挿抜を監視する管理ディスクの関係を一意に決定する。

【0037】LED203-0~203-nの代わりにそれぞれ接続調査手段207-0~207-nが設けられ、ディスク装置300-0~300-n内部にはLED制御手段301-0~301-nの代わりにそれぞれ接続状況記録手段302-0~302-nが設けられており、その接続形態は、図1と同じである。

【0038】接続調査手段207-0~207-nは、それぞれスイッチb208-0~208-nと接続しているかによって、ディスク装置300-0~300-nがディスクきょう体に格納されているかどうかを調べるために用いる。接続状況記録手段302-0~302-nは、ディスク装置300-0~300-nが抜き挿じられたことを記録しておくものである。

【0039】また、図5は第2の実施形態のRAID制御プログラム110をのブロック図である。図2との違いは、故障検出LED制御手段124が、活線挿抜制御手段125に代わったことである。活線挿抜制御手段125は、故障検出LED制御手段124と同様に、第1の実施形態と同じ方法でID管理テーブル130を作成し、管理情報保持手段111の中に保持する。また、第2の実施形態のID管理テーブル130は、故障したディスク装置300に対応する接続調査手段207を監視する管理IDを特定するためのものである。

【0040】次に、故障したディスク装置がユーザによって抜かれたことと、新しいディスク装置がユーザによって挿入されたことをホスト計算機100に通知する動作をしめす。活線挿抜制御手段125によるID管理テーブル130の作成、保持方法は、第1の実施形態と同様である。

【0041】あるディスク装置300が故障すると、ディスクインターフェース制御手段108を介して、ディスクインターフェース制御手段制御部123にどのディスク装置300が故障したかの通知が送られる。この情報をもとにして、RAID制御部122は、活線挿抜制御手段125に監視すべき接続調査手段207を問合わせる。活線挿抜制御手段125は、ID管理テーブル130をもとに故障したディスク装置300に関連する管理IDをRAID制御部122に知らせる。その管理IDのディスク装置300に対して、RAID制御部122は、ディスクインターフェース制御手段制御部123を介し、接続状況取得コマンドを発行する。

【0042】

【発明の効果】本発明によれば、RAIDの制御を行う手段とディスクきょう体を接続するのに特別な専用線を用いる必要がないので、低価格でディスクきょう体の管理を行うことができるという効果がある。

【0043】さらに、本発明によれば、ディスクインターフェースバス上に、ディスクきょう体管理専用デバイスを設ける必要がないので、ディスクインターフェース

バスに接続できるデバイス数を減らすことなく低価格でディスクきょう体の管理を行うことができるという効果がある。

【図面の簡単な説明】

【図1】第1の実施形態におけるシステム構成図である。

【図2】第1の実施形態におけるRAID制御プログラムのブロック図である。

【図3】第1の実施形態における管理情報保持手段のブロック図である。

【図4】第2の実施形態におけるシステム構成図である。

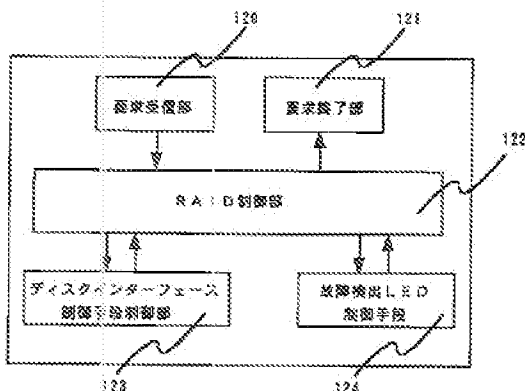
【図5】第2の実施形態におけるRAID制御プログラムのブロック図である。

【符号の説明】

100…ホスト計算機、101…ホストCPU、102…キャッシュ、103…チップセット、104…ローカルバス、105…システムバス、106…RAID制御手段、107…バスインターフェース、108…ディスクインターフェース制御手段、109…メインメモリ、110…RAID制御プログラム、111…管理情報保持手段、112…ユーザ領域、120…要求受信部、121…要求終了部、122…RAID制御部、123…ディスクインターフェース制御手段制御部、124…故障検出LED制御手段、125…活線挿抜制御手段、130…ID管理テーブル、200…ディスクきょう体、201…ディスクインターフェースバス、202…LED配線部、203-0～203-n…LED、204-0～204-n…スイッチa、205-0～205-n…バス接続部、206…LED管理ディスク決定手段、207-0～207-n…接続調査手段、208-0～208-n…スイッチb、209…接続調査配線部、210…活線挿抜管理ディスク決定手段、300-0～300-n…ディスク装置、301-0～301-n…LED制御手段、302-0～302-n…接続状況記録手段

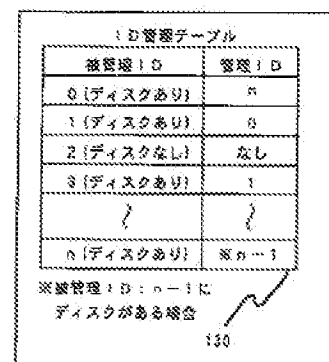
【図2】

図2

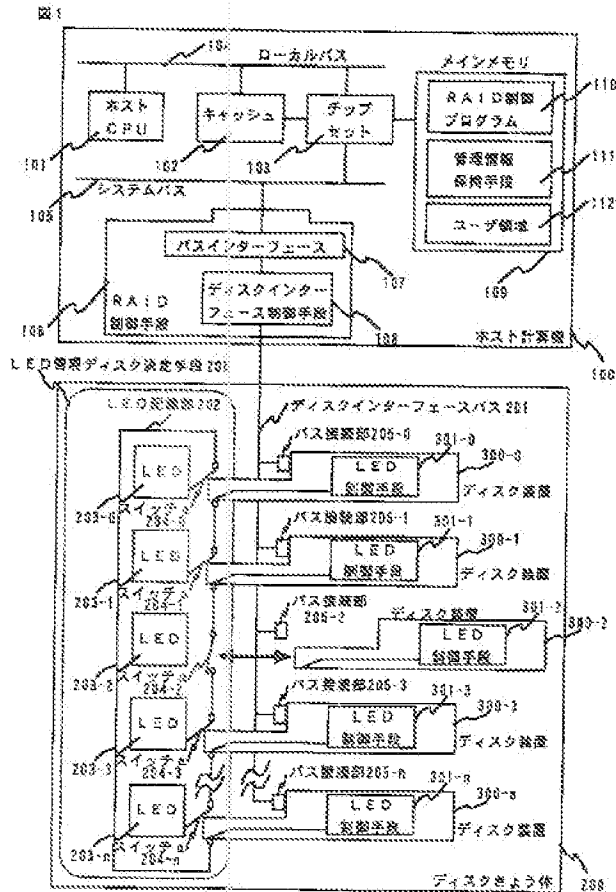


【図3】

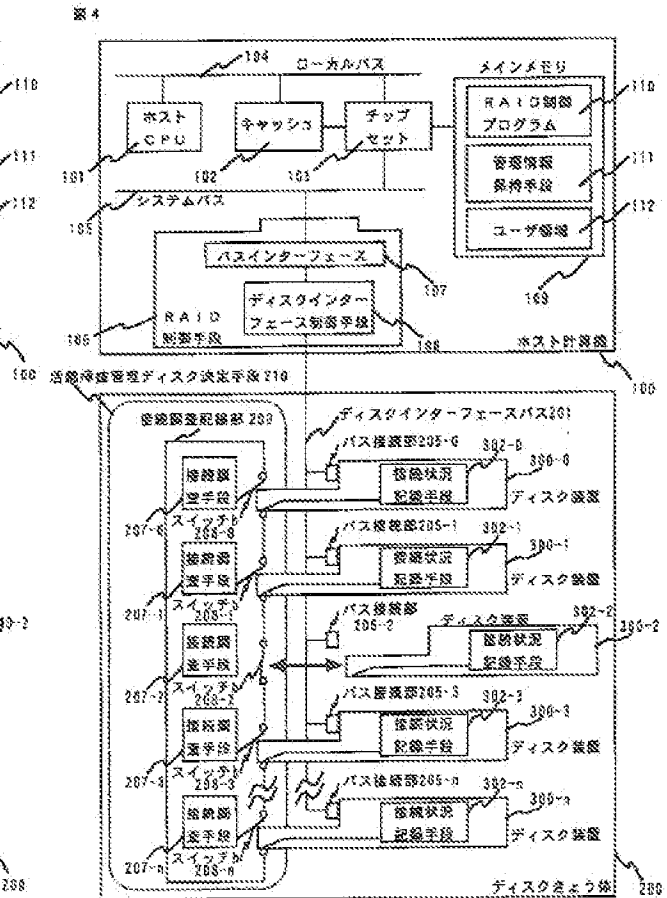
図3



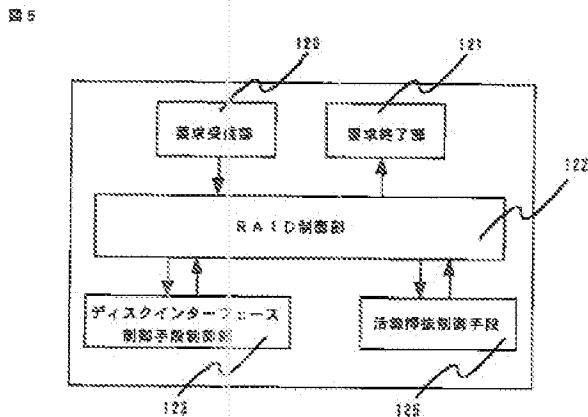
【図1】



【図4】



【図5】



フロントページの続き

(72)発明者 兼田 泰典

神奈川県川崎市麻生区王禅寺1099番地 株  
式会社日立製作所システム開発研究所内

(72)発明者 荒川 敦史

神奈川県川崎市麻生区王禅寺1099番地 株  
式会社日立製作所システム開発研究所内



(72)発明者 高野 隆弘  
神奈川県小田原市国府津2880番地 株式会  
社日立製作所ストレージシステム事業部内